# AN ENHANCED ELECTRONIC HEALTH COMMUNITY WITH KNOWLEDGE-BASED E-MAIL AND AGENT-BASED KNOWLEDGE SEARCH AND SHARING

Yu-N Cheah
*Health Informatics Research Group*
*School of Computer Sciences, Universiti Sains Malaysia, 11800 USM Penang, Penang, Malaysia*
*yncheah@cs.usm.my*

Faridah Abdul Rashid
*Department of Chemical Pathology*
*School of Medical Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia*
*faridah@kb.usm.my*

**ABSTRACT**

Electronic health communities have the potential to go beyond providing basic communication-type services such as online chat and discussion group, health directories and specialised portals for healthcare practitioners. This can be done by introducing artificial intelligence techniques and knowledge management processes, thus leading towards an Enhanced Electronic Health Community. In this paper, we propose a knowledge-enabled health community comprising of (1) an Intelligent Web Content Management application that not only provides regular communication services but also intuitive web publishing and personalised document search features, (2) a Knowledge-based E-mail application that allows e-mail to be shared and reused among healthcare practitioners for decision support, and (3) an Agent-based Knowledge Search and Sharing application that reduces the effort needed by healthcare practitioners to look for answers and to share useful material.

**KEYWORDS**

Health community, personalisation, knowledge management, e-mail.

## 1. INTRODUCTION

Electronic societies and communities are already a way of life in a number of domains, particularly in business and government. The electronic way of communicating with others around us within the same working environment and social context presents very convenient and effective means of sharing information and imparting knowledge. In the same way, electronic health communities are being established to facilitate better delivery of healthcare services to the masses.

In the context of healthcare, current electronic health communities focus on providing basic communication-type services to its members in the form of healthcare portals. Many such healthcare portals provide general information, general chat and discussion group services, health directories of healthcare practitioners and specialised 'portal within a portal' for medical practitioners.

These services are necessary and serve their purpose to a certain extent. However, we argue that electronic health communities have the potential to go beyond just making healthcare communication paperless or to go beyond encouraging the use of information and communication technologies in healthcare. We believe that users within electronic health communities, both the general public and healthcare practitioners, can be empowered by one another to use these technologies to make better decisions regarding their personal health and those of their patients.

With this aim in mind, we propose an Enhanced Electronic Health Community (EEHC) that capitalises on the basic functions of current electronic health communities or portals and enhances them by including

artificial intelligence techniques (e.g. constraint programming, natural language processing, etc.) and knowledge management processes (e.g. knowledge identification and knowledge sharing).

## 2. AN OVERVIEW OF THE ENHANCED ELECTRONIC HEALTH COMMUNITY

The EEHC consists of three main applications (see Figure 1):

1. *Intelligent Web Content Management*: This application is the main application of the EEHC. It provides the general public with relevant and personalised healthcare information such as those on healthy living, directory of healthcare practitioners, drugs and pharmacies, online chat and discussions, etc.
2. *Knowledge-based E-mail*: This application allows healthcare practitioners in the community to capitalise on potential knowledge stored in their e-mail messages via preexisting e-mail software. This application enhances the basic features found in the Intelligent Web Content Management application.
3. *Agent-based Knowledge Search and Sharing*: This application allows healthcare practitioners to search and share knowledge-rich material and documents with minimum effort via the use of intelligent agents. This application also enhances the Intelligent Web Content Management application.

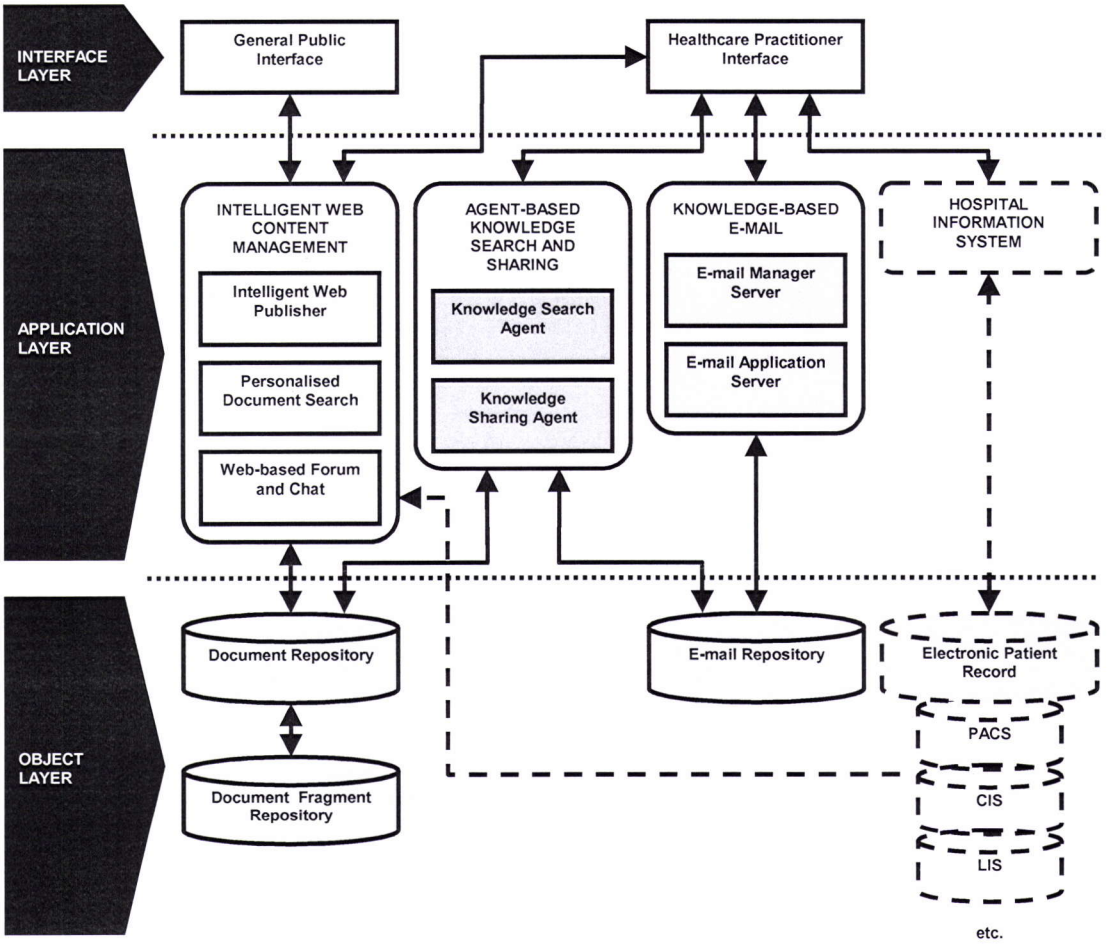


Figure 1. The Enhanced Electronic Health Community overall architecture

The three applications are organised into three conceptual layers:

1. *Object Layer*: This layer basically contains the knowledge-based e-mail repository, document and information repositories and electronic patient records (EPR). The EPR is potentially linked to other

repositories found in hospital-based systems such as Picture Archiving and Communication Systems (PACS), Clinical Information Systems (CIS), Laboratory Information Systems (LIS), etc.

2. *Application Layer*: The three main applications mentioned earlier reside in this layer. Another contributing application towards the EEHC is the Hospital Information System (HIS) that manages the EPR and all other hospital-based systems. However, we would not elaborate on the HIS as it is beyond the scope of this paper.
3. *Interface Layer*: This layer is essentially the client-side interfaces where healthcare practitioners and the general public would interact amongst themselves via the applications in the application layer.

# 3. INTELLIGENT WEB CONTENT MANAGEMENT

The Intelligent Web Content Management application can be viewed as a watered-down version of a generic multi-purpose portal that provides various content- and transaction-type services. It has three main components:

1. *Intelligent Web Publisher*: This web-publishing feature allows healthcare practitioners as well as the general public to upload and publish material on the web in an easy and organised way.
2. *Personalised Document Search*: This feature allows healthcare practitioners as well as the general public to not only search for documents stored in the repository (that was uploaded via the Intelligent Web Publisher) but also to receive personalised search results that matches the profile of the user.
3. *Web-based Forum and Chat*: This feature provides standard forum and chat services to the community to supplement regular communication, e.g. via telephone, e-mail and fax. As these services are somewhat standard, we would not be elaborating on this component.

## 3.1 Intelligent Web Publisher

In any health community, dissemination of information is important and a mechanism to allow easy publication and posting of information, similar to that of a bulletin board to post articles, announcements and directories is vital. The paper-based bulletin board may be simple enough to maintain but in an electronic health community, publication of online web-based material poses some inconvenience, especially with the need to be familiar with HTML, web-servers, etc. With the Intelligent Web Publisher, we aim to make web publishing convenient enough for use within a health community.

In implementing the Intelligent Web Publisher component, five main sub-components are involved (see Figure 2):

1. *Web Publisher Interface*: The healthcare practitioner or a member of the public would basically interact with an intelligent Web Publisher Interface that accepts the materials to be published. This interface allows the user to browse through the directories or folders in the source-computer and to select the file. The user is also required to include keywords and descriptions on the content of the file to be published and to decide on the appropriate sub-domain category for the file.
2. *Format Converter*: After the material has been submitted via the Web Publisher Interface, the Format Converter intelligently detects the format of the file. We feel that the detection of the format via the file extension alone is inaccurate at times because file extensions can be easily modified for one reason or another. Following this, the Format Converter employs an algorithm to convert the file into HTML. For example, if a healthcare practitioner would like to upload a MS-Word document about diabetes for the general public to read, the document is converted intuitively into HTML without having the need for the healthcare practitioner to be concerned about formatting consistencies, e.g. font sizes, headings, etc.
3. *Document Categoriser and Verifier*: Once the file has been converted, it is then analysed to verify whether the content of the file relates to the keywords, descriptions and category provided by the user via the Web Publisher Interface. This analysis may be carried out using latent semantic analysis (Landauer et al, 1998) or keyword extraction (Matsuo and Ishizuka, 2003) techniques. This process ensures that the keywords, descriptions and category are not misleading or inaccurate.
4. *Publishing Linker*: The verified file is then linked to a referring page under the appropriate category via automatically generated hyperlinks. This updated referring page is then made available on the health community's main webpage.

5. *Document Repository*: All published material are stored in a document repository. This repository also serves as a source for information and would be searched by the Personalised Document Search component.

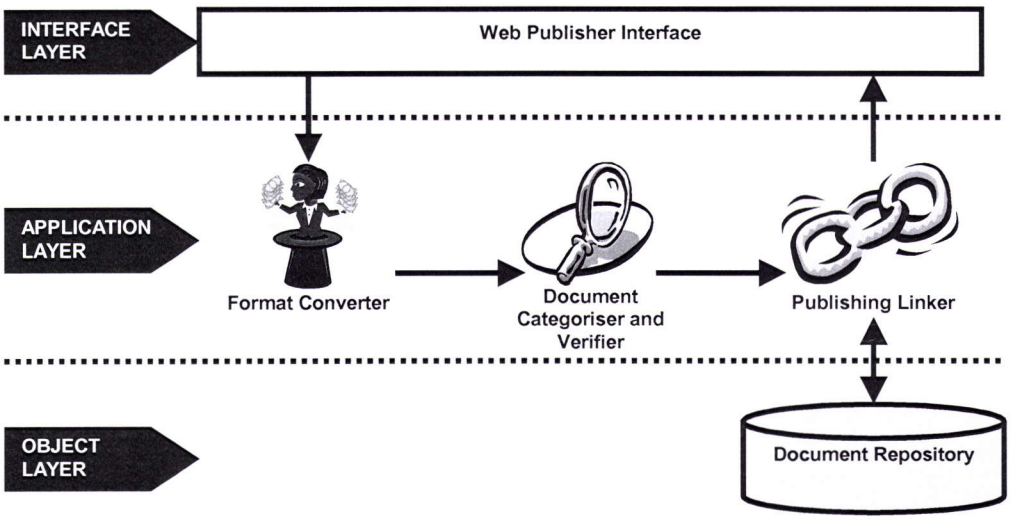


Figure 2. Intelligent Web Publisher architecture

## 3.2 Personalised Document Search

The search for information by healthcare practitioners and the general public usually results in generic search results that may be applicable to a wide audience. While this may be useful in some cases, many users, especially those in healthcare, may require specific search results that are relevant to their needs. These needs may be different along two lines:

- *Different level of expertise of the user*: Healthcare practitioners and the general public would have different level of knowledge on a particular disease or treatment.
- *Different personal or family health history of user*: An adult member of the general public who has diabetes may not be interested to see search results on juvenile diabetes while information about complications of diabetes in middle- or old-age may be more relevant.

Therefore, the Personalised Document Search component takes into account the profile of the user who is requesting for a particular piece of information in order to personalise the results. Efforts to personalise documents are not new. Some of these efforts require a very structured document base (Abidi and Chong, 2001). Our proposed Personalised Document Search component hopes to produce personalised documents that reads more naturally. The following are sub-components of the Personalised Document Search component (see Figure 3):

1. *Document Fragment Generator*: This sub-component capitalises on the content of the document repository of the Intelligent Web Publisher. These categorised and verified documents are further analysed and broken down into shorter fragments based on its content and topic. The analysis can be done by detecting section headings and employing latent semantic analysis and keyword extraction techniques mentioned earlier to create document fragments. These document fragments are similar to the topic-specific documents developed by (Abidi and Chong, 2001). The document fragments proposed here essentially follow the same idea except that they are less structured that the topic-specific documents, i.e. they would preserve much of the paragraph and sentence structures but would be accompanied by essential meta-data such as keywords, topic and constraints.
2. *Personalised Search Interface*: The search interface accepts a user's search request and also presents a questionnaire to the user to find out more regarding the user's latest health status and lifestyle, i.e. dietary habits and physical activities.
3. *Intelligent Document Composer*: This sub-component carries out the actual personalisation of the search documents. It takes into account the user's electronic patient record and inputs from the questionnaire to

create a user's current health profile. Essentially, the Intelligent Document Composer selects the various document fragments which contain the relevant meta-data that matches the user's current health profile and the user's latest health status and lifestyle. The document fragments are then combined to form a complete document.

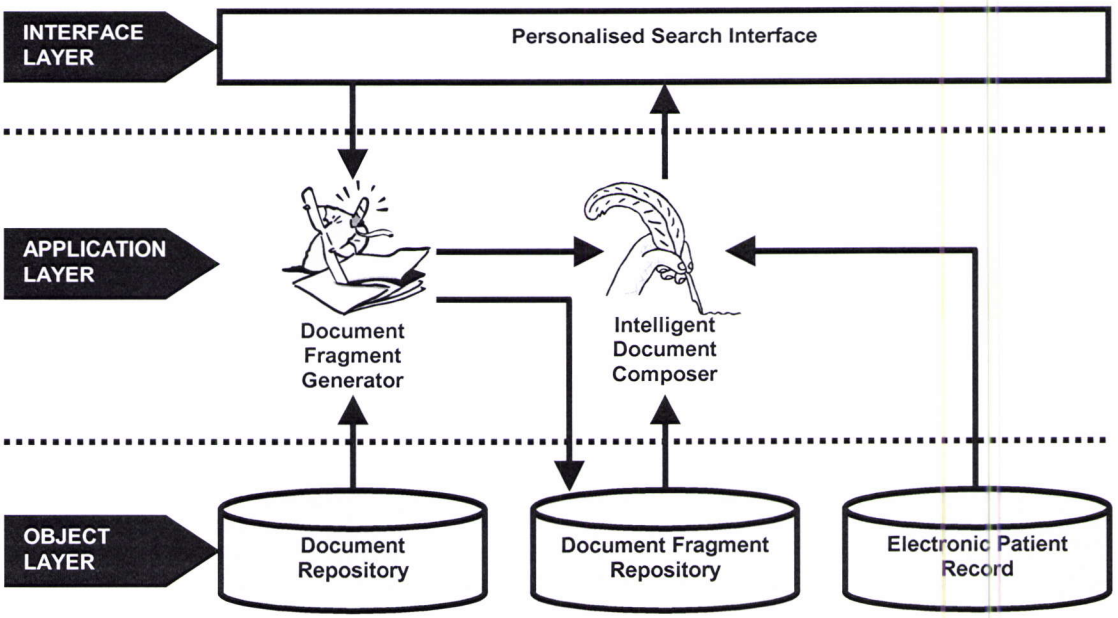


Figure 3. Personalised Document Search component architecture

## 4. KNOWLEDGE-BASED E-MAIL

E-mail has revolutionised the way individuals in a community interact. Perhaps, for the general public in a health community, it is sufficient to use regular e-mail systems for communication with their doctors, pharmacists and support groups. However, among healthcare practitioners, e-mail could prove to be a vast repository of experience and knowledge that could be reused and re-visited rather than just be archived in the e-mail folders of one's computer. The key feature of the Knowledge-based E-mail is the e-mail evaluation mechanism which allows the 'micro-community' of healthcare practitioners to better decide which e-mail messages contain useful knowledge and experience that deserve a 'second-look'.

The Knowledge-based E-mail consists of two main components (see Figure 4):

1. *E-mail Manager Server*: The manager server supports the basic functions of sending and receiving e-mail messages from various e-mail clients used by the community of healthcare practitioners. This server would employ alternative protocols for handling e-mail messages unlike traditional e-mail protocols that popularly utilise POP and SMTP protocols (Cheah and Lim, 2003).
2. *E-mail Application server*: The application server provides the value-added functions of the Knowledge-based E-mail that the user interacts with.

### 4.1 E-mail Manager Server

The main sub-components of the E-mail Manager Server are as follows:

- *Intelligent E-mail Manager*: This component handles all incoming e-mail transactions and decides on the next course of action based on who the sender or recipients are, the e-mail subject or even the e-mail size.

- *Mailet*: This mailet component receives instructions from the Intelligent E-mail Manager. Possible actions include the creation of discussion groups for a particular healthcare issue/problem or the forwarding of e-mail messagess to the intended group of recipients.
- *E-mail Formatter*: This component works in tandem with the Mailet to perform parsing functions to capture relevant message content (e.g. sender, recipient, date, time, etc.) as well as to reformat e-mail messages before they are forwarded to recipients or before it is stored in the repository.

## 4.2 E-mail Application Server

The second component is the application server which consists of the following sub-components:

- *Service Manager*: This component receives service requests from users. This can be in the form of registration requests, etc.
- *Evaluation Engine*: This engine allows a recipient to evaluate an e-mail's content (this may be a reply to a query/problem) in terms of its quality (i.e. usefulness, relevance, accuracy, etc.). The evaluation details are stored together with the e-mail concerned and would facilitate searches for quality answers from the e-mail content.

Figure 4 illustrates the Knowledge-based E-mail application architecture.

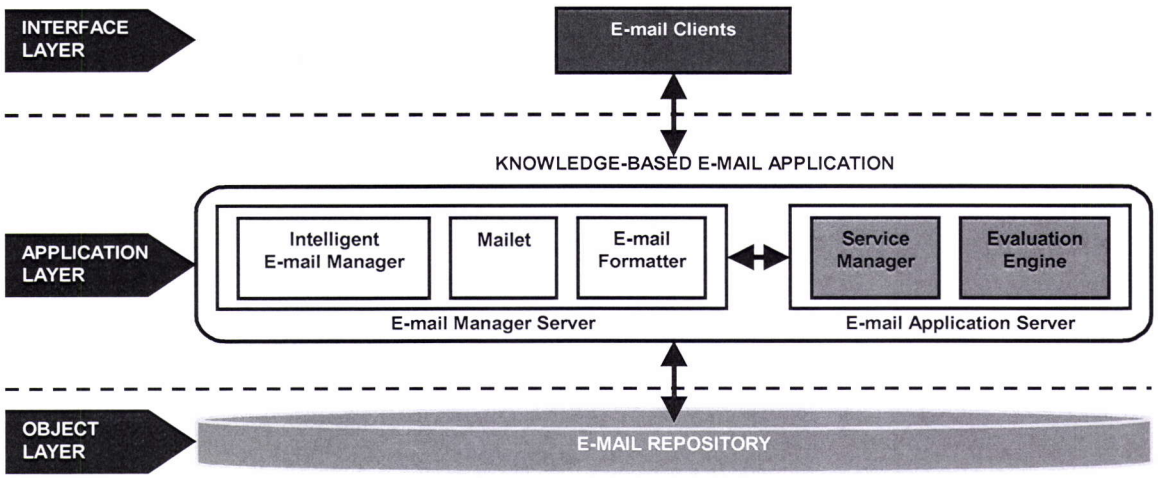


Figure 4. The Knowledge-based E-mail architecture (Cheah and Lim, 2003)

# 5. AGENT-BASED KNOWLEDGE SEARCH AND SHARING

While the Personalised Document Search component of the Intelligent Web Content Management application may be able to cater for most information request by, both, healthcare practitioners and the general public, healthcare practitioners may have the need for specialised knowledge that cannot be obtained from the document repository of the Personalised Document Search component. In these cases, the community of healthcare practitioners need to be able to search and share specialised knowledge amongst themselves.

The Agent-based Knowledge Search and Sharing application is built on a knowledge management framework that provides the basic and generic architecture of all intelligent agents components including communication conventions, etc. (Lacher and Koch, 2000). These agents are then defined further to address the different tasks assigned to it. The Agent-based Knowledge Search and Sharing application consists of two main agent components (Cheah et al, 2004):

1. *Knowledge Search Agent*: This agent basically allows users to search for knowledge without having to be 'too proactive', i.e. without having to deal with the intricacies of using a search engine or having to ask someone in the health community.

2. *Knowledge Sharing Agent*: This agent complements the efforts of the search agent by allowing users to publicise available knowledge, again, without having to be 'too proactive' in view that time is precious for most healthcare practitioners.

Figure 5 illustrates the architecture of the Agent-based Knowledge Search and Sharing application.

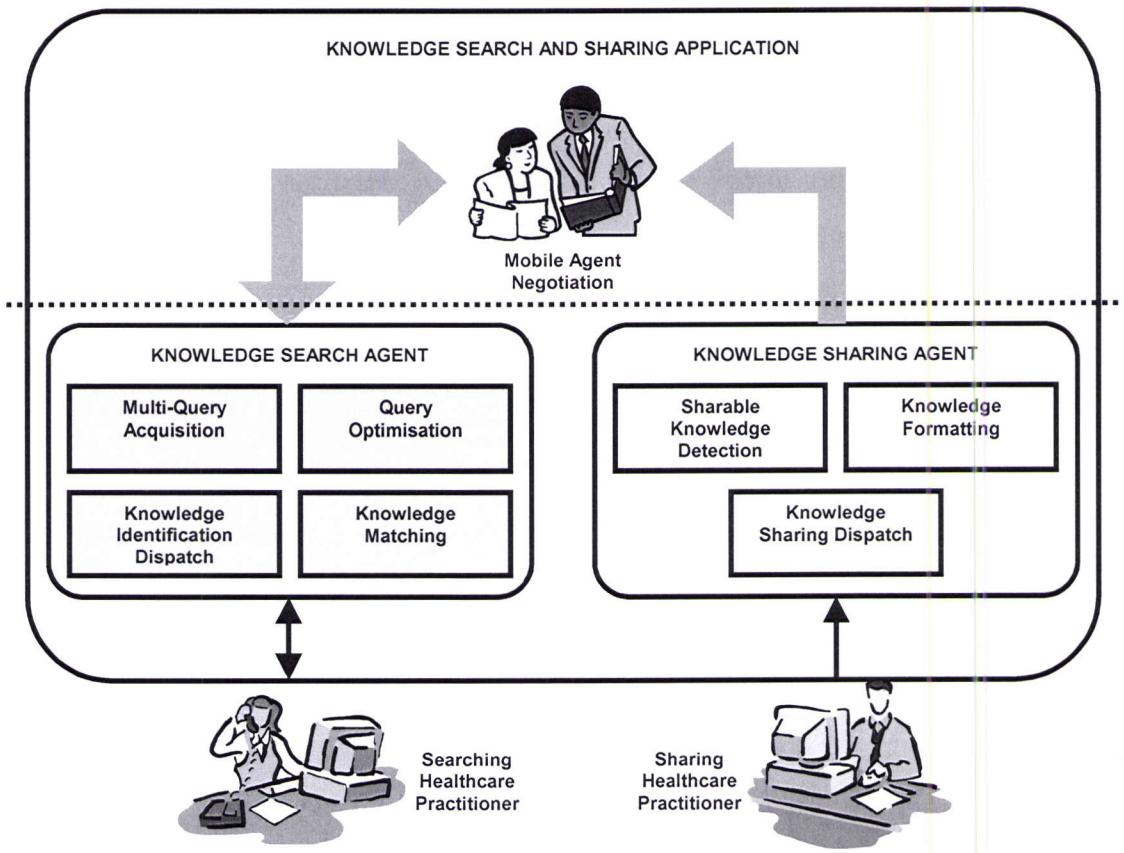


Figure 5. The Agent-based Knowledge Search and Sharing application architecture

## 5.1 Knowledge Search Agent

The Knowledge Search Agent facilitates healthcare practitioners to go about looking for knowledge. These search protocols allows search criteria to be specified more intuitively and explicitly to facilitate current search algorithms to produce more accurate search results. Main sub-components of this search agent are:

- *Multi-Query Acquisition*: This sub-component intelligently captures the queries of the users using a novel protocol that dynamically adapts to the way the user specifies a query. This minimises any ambiguity in the query.
- *Query Optimisation*: Upon receiving the query, this sub-component reformats the query into a concise and easily transmitted format.
- *Knowledge Identification Dispatch*: This sub-component takes on the behaviour of a mobile agent to 'publicise' the query to other mobile agents, i.e. to other search agents who may have similar search requests so as to pool their efforts.
- *Knowledge Matching*: This sub-component ultimately identifies (searches) and retrieves the required knowledge for the user.

## 5.2 Knowledge Sharing Agent

The Knowledge Sharing Agent complements the efforts of the knowledge search agents where healthcare practitioners can share relevant information and experience (or knowledge) seamlessly, i.e. via commonly used applications, e.g. e-mail, word-processing, file folder browsers, etc. Main sub-components of the Knowledge Sharing Agent are:

- *Sharable Knowledge Detection*: This sub-component employs natural language processing to proactively detect documents or e-mail messages that potentially contain experience-related material and prompt the user accordingly. Alternatively, it also allows reactive submission of such documents by the healthcare practitioners themselves.
- *Knowledge Formatting*: This sub-component complements the Knowledge Matching sub-component by ensuring that the shared knowledge is in a format that can be effectively identified (or searched) and retrieved.
- *Knowledge Sharing Dispatch*: This sub-component is a mobile agent that 'publicises' sharable knowledge.

This application potentially acts as a secondary feed to the Personalised Document Search component's repository.

# 6. CONCLUSION

The Enhanced Electronic Health Community is presently at the proposal stage. At the component level, prototypes of an online document personalisation system (Abidi and Chong, 2001) and a peer-to-peer document search and sharing system (i.e. not knowledge structures) is presently available. We are progressing towards having the Intelligent Web Publisher and the Knowledge-based E-mail components implemented.

With the ever increasing use of the Internet by healthcare practitioners and the general public on healthcare matters (Provost et al, 2003), we are confident that they would welcome value-added features, such as those proposed in this paper, that would enhance their interaction in an electronic health community.

# REFERENCES


Abidi, S.S.R. and Chong, Y.H., 2001. An Intelligent Info-Structure for Composing and Publishing Personalised Healthcare Information Over the Internet. *Proceedings of the 14th IEEE Symposium on Computer Based Medical Systems (CBMS 2001)*. Bethesda, USA.

Cheah, Y.-N. and Lim, K.G., 2003. KM-Mail: An Approach for Sharing, Evaluating and Organizing Knowledge-Rich E-mail, *The 2003 International Conference on Information and Knowledge Engineering (IKE 2003)*. Las Vegas, Nevada, USA.

Cheah, Y.-N. et al, 2004. A Platform for Enterprise-Wide Healthcare Knowledge Management. *Proceedings of the Knowledge Management International Conference and Exhibition (KMICE 2004)*, Penang, Malaysia.

Matsuo, Y. and Ishizuka, M., 2003. Keyword Extraction from a Single Document using Word Co-occurrence Statistical Information. *Proceedings of the 16th International Florida Artificial Intelligence Research Society Conference (FLAIRS 2003)*. Florida, USA, pp. 392-396.

Lacher, M.S. and Koch, M., 2000. An Agent-based Knowledge Management Framework. *Proceedings of the AAAI Spring Symposium*. Stanford, California, USA, pp. 145-147.

Landauer, T.K. et al, 1998. Introduction to Latent Semantic Analysis. *Discourse Processes*, Vol. 25, pp. 259-284.

Provost, M. et al, 2003. Opinions and e-Health Behaviours of Patients and Health Professionals in the U.S.A. and Europe. *Proceedings of Medical Informatics Europe (MIE 2003)*. St. Malo, France, pp. 695-700.

## Call for Papers

The IADIS e-Society 2004 conference aims to address the main issues of concern within the Information Society. This conference covers both the technical as well as the non-technical aspects of the Information Society. Broad areas of interest are eGovernment / eGovernance, eBusiness / eCommerce, eLearning, eHealth, Information Systems, and Information Management. These broad areas are divided into more detailed areas (see below). However innovative contributes that don't fit into these areas will also be considered since they might be of benefit to conference attendees.

Topics related to e-Society are of interest. These include best practice, case studies, strategies and tendencies in the following areas:

### eGovernment/eGovernance

May include issues relating to:
· Accessibility
· Democracy and the Citizen
· Digital Economies
· Digital Regions
· eAdministration
· eGovernment Management
· eProcurement
· Global Trends
· National and International Economies
· Social Inclusion

### eBusiness / eCommerce

May include issues relating to:
· Business Ontologies and Models
· Digital Goods and Services
· eBusiness Models
· eCommerce Application Fields
· eCommerce Economics
· eCommerce Services
· Electronic Service Delivery
· eMarketing
· Languages for Describing Goods and Services
· Online Auctions and Technologies
· Virtual Organisations and Teleworking

### eLearning

May include issues relating to:
· Collaborative Learning
· Curriculum Content Design & Development
· Delivery Systems and Environments
· Educational Systems Design
· eLearning Organisational Issues
· Evaluation and Assessment
· Virtual Learning Environments and Issues
· Web-based Learning Communities

### eHealth

May include issues relating to:
· Data Security Issues
· eHealth Policy and Practice
· eHealthcare Strategies and Provision
· Legal Issues
· Medical Research Ethics
· Patient Privacy and Confidentiality

**Information Systems**

May include issues relating to:
· Electronic Data Interchange (EDI)
· Intelligent Agents
· Intelligent Systems
· IS Security Issues
· Mobile Applications
· Multimedia Applications
· Payment Systems
· Protocols and Standards
· Software Requirements and IS Architectures
· Storage Issues
· Strategies and Tendencies
· System Architectures
· Telework Technologies
· Ubiquitous Computing
· Virtual Reality
· Wireless Communications

**Information Management**

May include issues relating to:
· Computer-Mediated Communication
· Content Development
· Cyber law and Intellectual Property
· Data Mining
· ePublishing and Digital Libraries
· Human Computer Interaction
· Information Search and Retrieval
· Knowledge Management
· Policy Issues
· Privacy Issues
· Social and Organizational Aspects
· Virtual Communities
· XML and Other Extensible Languages

The Conference will be composed of several types of contributions:

- **Full Papers** – These include mainly accomplished research results and have 8 pages at the maximum (5,000 words).
- **Short Papers** – These are mostly composed of work in progress reports or fresh developments and have 4 pages at maximum (2,500 words).
- Posters / Demonstrations – These have one page at the maximum (625 words) besides the poster itself (or demonstration) that will be exposed at the conference.
- **Tutorials** – Tutorials can be proposed by scholars or company representatives. A proposal of maximum 250 words is expected.
- **Panels** – Discussions on selected topics will be held. A proposal of maximum 250 words is expected.
- **Invited Talks** – These will be made of contributions from well-known scholars and company representatives. An abstract will be included in the conference proceedings.
- **Doctoral Consortium** - A Doctoral Consortium will discuss in group, individual projects and on going work of PhD students. Prospective students should send a report of their PhD projects and work so far with a maximum of 4 pages (2,500 words).
- **Corporate Showcases & Exhibitions** – The former enables Companies to present recent developments and applications, inform a large and qualified audience of your future directions and showcase company's noteworthy products and services. There will be a time slot for companies to make their presentation in a room. The latter enables companies the opportunity to display its latest offerings of hardware, software, tools, services and books, through an exhibit booth. For further details please contact the publicity chair - es2004@iadis.org.

**All submissions, except invited talks, are subject to a blind refereeing process.**
**Important Dates:**
- Submission Deadline: ~~26 March 2004~~ 3 May 2004 (Special Call - Short papers, posters and doctoral papers only)
- Notification to Authors: 26 April 2004
- Final Camera-Ready Submission and Early Registration: Until 14 May 2004
- Late Registration: After 14 May 2004
- Conference: Spain, 16 to 19 July 2004